



## OPEN ACCESS

## EDITED BY

Guang-Long Wang,  
Huaiyin Institute of Technology, China

## REVIEWED BY

Aureliano Bombarely,  
Polytechnic University of Valencia, Spain  
Kang Zhang,  
Chinese Academy of Agricultural Sciences,  
China

## \*CORRESPONDENCE

Rainer Schwacke  
✉ r.schwacke@fz-juelich.de

RECEIVED 31 March 2025

ACCEPTED 02 June 2025

PUBLISHED 24 June 2025

## CITATION

Schwacke R, Bolger ME and Usadel B (2025)  
PubPlant – a continuously updated  
online resource for sequenced  
and published plant genomes.  
*Front. Plant Sci.* 16:1603547.  
doi: 10.3389/fpls.2025.1603547

## COPYRIGHT

© 2025 Schwacke, Bolger and Usadel. This is  
an open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# PubPlant – a continuously updated online resource for sequenced and published plant genomes

Rainer Schwacke<sup>1\*</sup>, Marie E. Bolger<sup>1</sup> and Björn Usadel<sup>1,2</sup>

<sup>1</sup>Forschungszentrum Jülich, Institute of Bio- and Geosciences, IBG-4 Bioinformatics, CEPLAS, BioSC, Jülich, Germany, <sup>2</sup>Heinrich-Heine University Düsseldorf, Faculty of Mathematical and Natural Sciences, Institute for Biological Data Science, CEPLAS, Düsseldorf, Germany

Advances in next-generation sequencing technologies over the last decade have substantially reduced the cost and effort required to sequence plant genomes. Whereas early efforts focused primarily on economically important crops and model species, attention has now turned to a broader range of plants, including those with larger and more complex genomes. In 2024, the genomes of 500 plant species were published, including 370 sequenced for the first time. Tracking and providing access to published plant genomes (now covering more than 1800 species) is an invaluable service for plant researchers. PubPlant is an online resource that serves this purpose by cataloging published plant genome sequences and offering multiple visualizations ([https://www.plabipd.de/pubplant\\_main.html](https://www.plabipd.de/pubplant_main.html)). It includes a chronology of genome publications, and cladograms to display the phylogenetic relationships among the sequenced plants. An overview diagram for seed plants highlights taxonomic orders and families with sequenced species and reveals those that have been overlooked thus far. As a use case for PubPlant, we evaluated the status of sequenced food crops. We found that the five plant families featuring the most food crops were those containing the most sequenced plant species.

## KEYWORDS

sequenced plant genomes, plant genome publication, timeline, cladogram, Archaeplastida

## Introduction

The first plant genome to be sequenced was that of the laboratory model *Arabidopsis thaliana*, published in the year 2000 ([The Arabidopsis Genome Initiative, 2000](#)). It took 10 further years to achieve the milestone of 20 sequenced plant genomes, but only another 4 years to pass 100 genomes, and by the year 2020 the milestone of 500 plant genomes had been

achieved. Remarkably, 500 additional plant genomes were sequenced in the next 2 years (one tenth of the time needed for the first 500). This progress is still accelerating, mainly due to the advent of third-generation long-read sequencing technologies (Jiao and Schneeberger, 2017) and their continual refinement (Dumschott et al., 2020; Pucker et al., 2022). Advances in sequencing technologies have gone hand in hand with the development of more powerful bioinformatics algorithms for the assembly and annotation of genomic data. Genome assembly tools such as hifiasm (Cheng et al., 2021a) and verkko (Rautiainen et al., 2023) can integrate data from the two most popular long-read sequencing technologies, namely nanopore sequencing developed by Oxford Nanopore Technologies and single-molecule real-time sequencing commercialized as the PacBio platform by Pacific Biosciences (van Rengs et al., 2022).

The advances in third-generation sequencing have reduced the cost and effort needed for genome sequencing to such a degree that it is now within the means of even moderately-funded research groups. It is therefore unsurprising that recent years have seen a dramatic increase in the number of plant whole-genome sequencing projects (Figure 1). By the end of 2024, more than 1800 plant species had been sequenced, more than 500 of which have been sequenced twice and more than 200 of which have been sequenced three or more times. The number of plant species that have been sequenced is increasing steadily, but the number of individual sequenced genomes is increasing at a much faster rate due to the re-sequencing of the same species multiple times. Re-sequencing is driven by the repetition of earlier sequencing efforts using more advanced technologies to obtain better and more complete genomes as well as intra-species pan-genome projects involving the sequencing of multiple individuals (such as different varieties,

cultivars or ecotypes) within a species to explore the full genomic landscape (Golitz et al., 2016). Prominent examples of staple crop pan-genomes include maize (Hufford et al., 2021), barley (Jayakodi et al., 2020), wheat (Jiao et al., 2024), rice (Qin et al., 2021) and potato (Tang et al., 2022; Bozan et al., 2023), as well as fruit crops such as tomato (Gao et al., 2019; Zhou et al., 2022), and beverage crops such as tea (Chen et al., 2023; Tariq et al., 2024).

The history of plant genome sequencing has been summarized at intervals to highlight the status of sequenced plant genomes at the time of publication, often focusing on particular technological advances (Michael and Jackson, 2013; Chen et al., 2019; Kersey, 2019; Shirasawa et al., 2021; Kress et al., 2022; Sun et al., 2022; Bernal-Gallardo and de Folter, 2024). But such is the pace of change that such review articles are often out of date by the time they are published. One attempt to present a more frequently updated resource is the Plants-Genomes-Technologies (N3) database (Xie et al., 2024). The current iteration is version 3.0 (accessed February 17, 2025), which was published on January 11, 2024. Here, we describe an additional online resource called PubPlant ([https://www.plabipd.de/pubplant\\_main.html](https://www.plabipd.de/pubplant_main.html)) that has tracked and continuously updated published plant genome sequences for almost a decade.

## Methods

### Tracking published plant genomes

PubPlant tracks published plant genomes using manual search and curation methods. Searches are conducted using the Google search engine, cited reference searches in the Web of Science

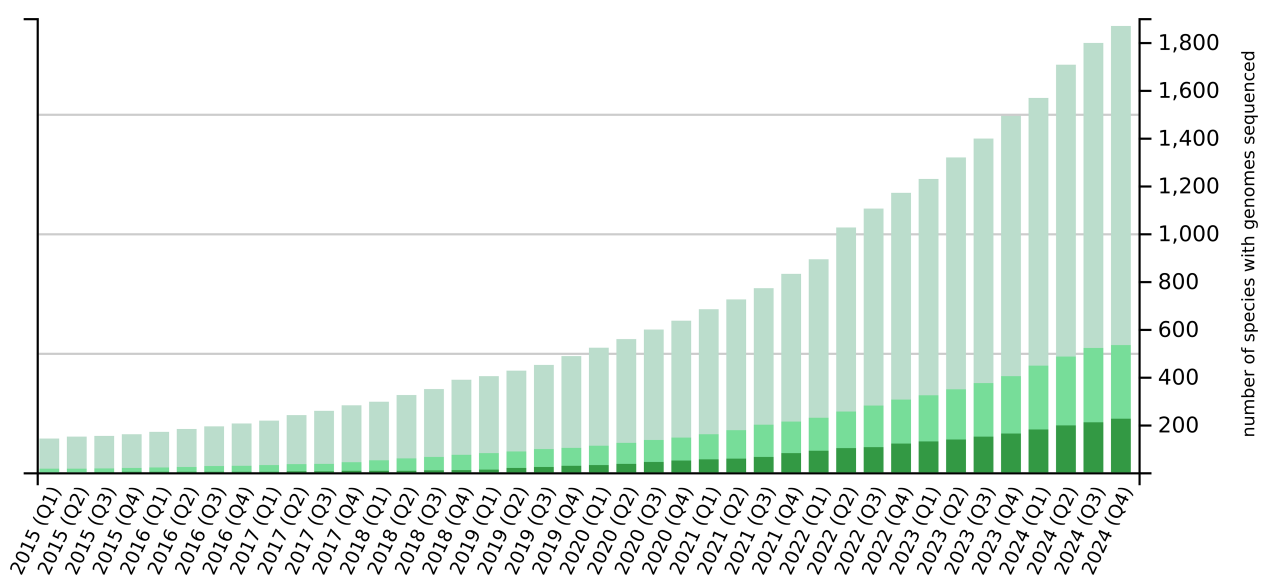


FIGURE 1

Number of plant species with sequenced and published genomes over time. The light green bars show the number of species that have been sequenced at least once, the medium green bars represent species that have been sequenced at least twice, and the dark green bars are those that have been sequenced three or more times. The bar chart shows quarterly data for the past decade.

database (Clarivate, 2025), and the NCBI PubMed citation database (PubMed, 2025). All search results are reviewed manually, and relevant information is extracted and entered into JavaScript Object Notation (JSON) files via copy and paste. These manually curated JSON files serve as the data source for the web-based diagrams that visualize the current status of sequenced plant genomes. The JSON source files also encode phylogenetic information, with taxonomic relationships based on the classifications listed in Table 1. Diagrams are generated client-side using a JavaScript script that employs the D3.js library (v4, <https://d3js.org>) for cladogram rendering, and the Vis.js library (v3.1, <https://visjs.org>) for the timeline visualizations.

To be included in PubPlant, the sequenced genome must belong to a plant (from the Archaeplastida group, which includes land plants, charophytes, green algae, glaucophytes, red algae and Rhodelphidophyta) and must be comprehensively described in a peer-reviewed journal. This means that the reads must be assembled into contigs, the contigs must have undergone scaffolding, and structural gene annotation must be completed. Accordingly,

TABLE 1 Resources used for phylogenetic classification in the diagrams presented by PubPlant.

Plant group	Classification
Angiosperms	APG-IV system (The Angiosperm Phylogeny Group, 2016; Zuntini et al., 2024)
Gymnosperms	Yang et al. (2022)
Lycophytes and ferns	PPG-I system (The Pteridophyte Phylogeny Group, 2016; Hassler, 2025)
Bryophytes	Bechteler et al. (2023)
Primary algae	Guiry and Guiry (2025)
Others	NCBI Taxonomy (Schoch et al., 2020); COL (2025)

PubPlant excludes highly fragmented genomes or those with incomplete or missing structural annotations.

If a publication includes the genomes of multiple accessions representing a single species, such as an intra-species pan-genome, PubPlant counts the species as sequenced and published once. If a pan-genome is based on the genomes of different species – often described as a super pan-genome (Khan et al., 2020) – each species is counted individually. The accepted scientific names of vascular plant species are cross-checked against Plants of the World Online (POWO), 2025 (Govaerts et al., 2021). The lowest recognized taxonomic rank is the species. Accordingly, genomes of different subspecies or varieties are pooled.

Food crops

All statistical data concerning major food crop species were extracted from the Food and Agriculture Organization Corporate Statistical Database (FAOSTAT, 2024). The data provided by FAOSTAT refer to crop categories with an annual production of more than 5000 tons in 2022. Each of these categories contains one or several food crops. The NCBI Taxonomy Database (Schoch et al., 2020) was used to assign correct scientific names to individual crops.

Results and discussion

Chronology of published plant genomes

PubPlant’s timeline view displays sequenced plant genomes according to the chronology of their first publication dates (Figure 2, a full list of published plant genomes is presented in Supplementary Table 1). If there are two publications for the same

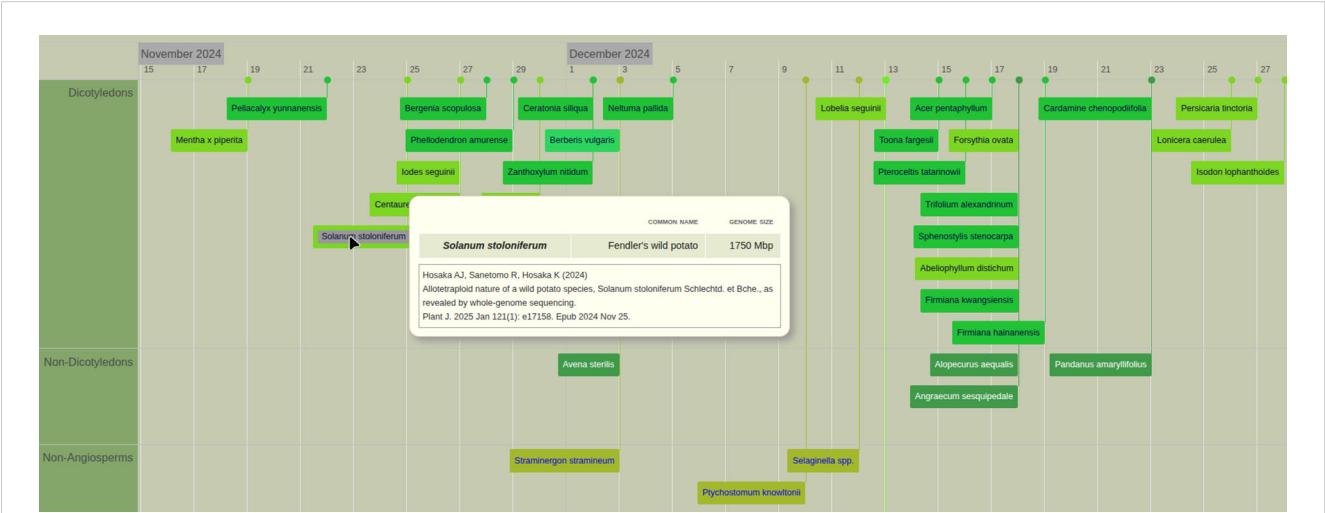
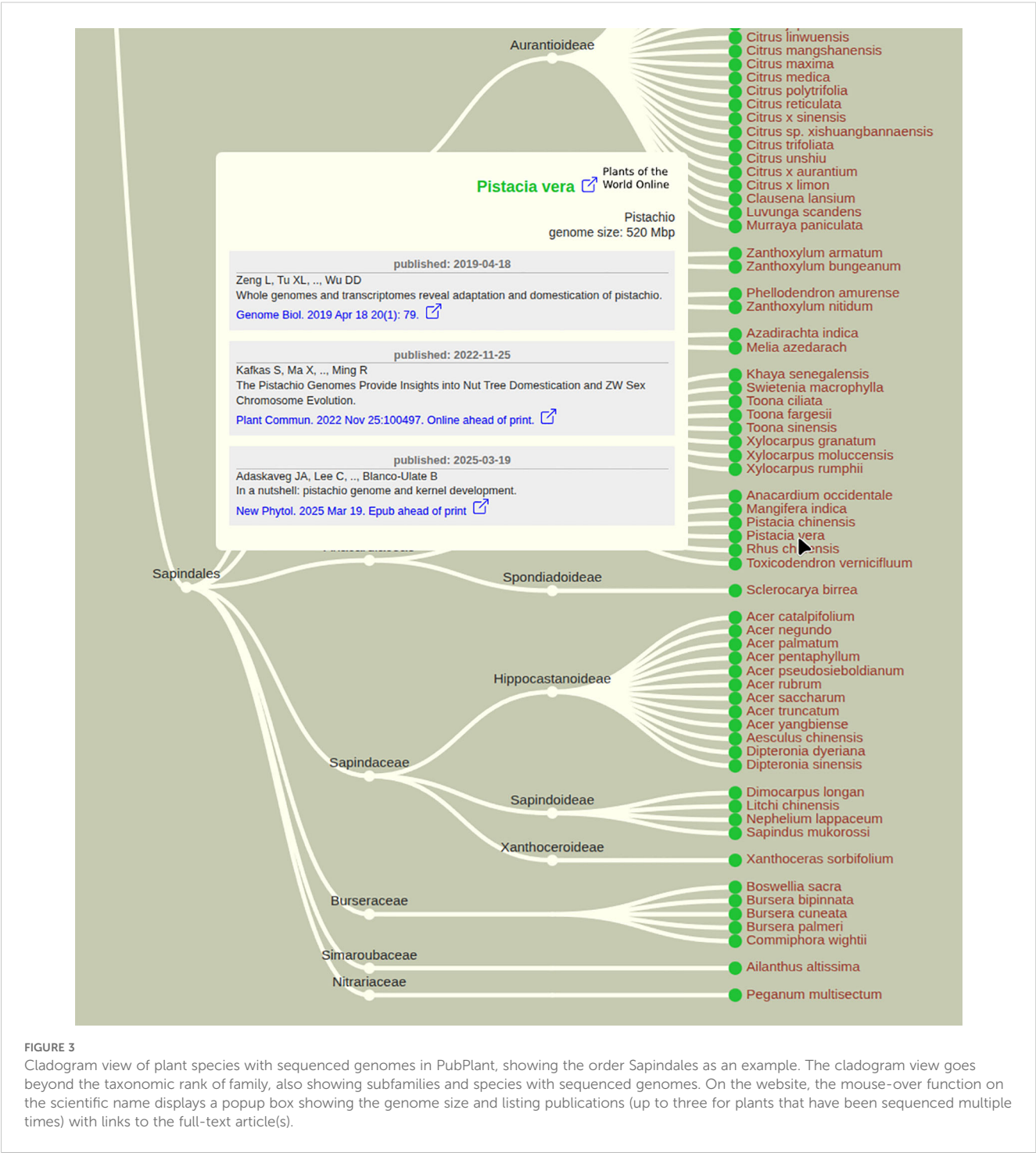


FIGURE 2 Timeline view of published plant genomes in PubPlant. The colored boxes refer to sequenced genomes of an individual species or a genus and multiple species (spp. for *species pluralis*) if the publication describes several species of the same genus. On the website, the mouse-over function displays the scientific name (or a list of scientific names if the box refers to several species of the same genus, as in the example shown for *Selaginella* spp.), the common plant name, the genome size, and the citation of the first publication. Clicking on the box links to the full-text publication in a new window.

plant species on the same date, the publication with the earliest acceptance date is listed. The timeline view groups the plant entries as belonging to dicotyledons, non-dicotyledons, non-angiosperms or algae (Figure 2).

Phylogeny of sequenced plant genomes

PubPlant’s cladogram view arranges published plant genomes according to the phylogenetic position of each species (Figure 3).



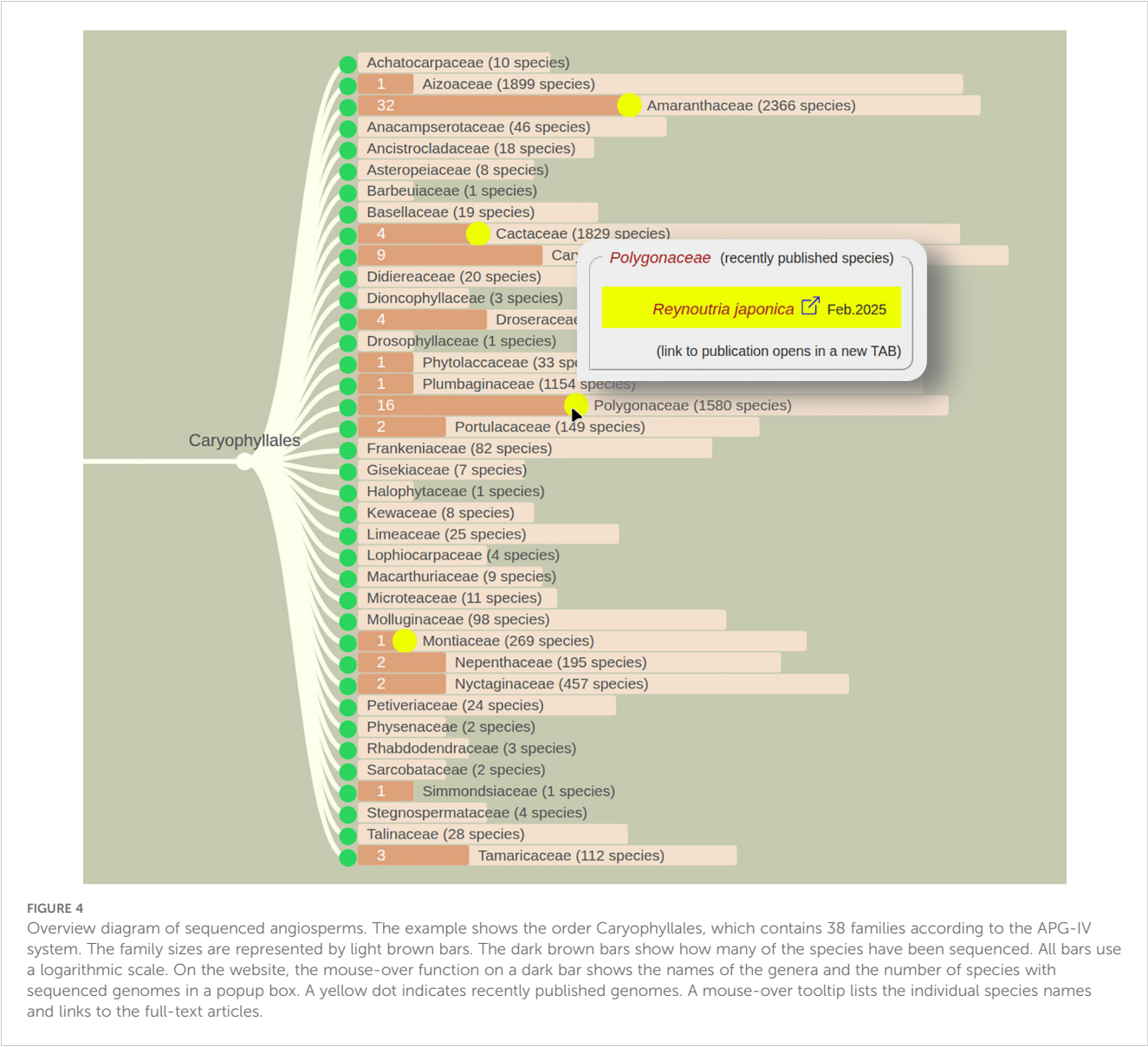
The cladograms for flowering and non-flowering plants are displayed separately to improve legibility. Each entry in the cladogram provides a mouse-over function that displays a popup box containing the scientific name, common name and genome size of the plant. Publication details (one or more publications), including links to the corresponding full-text articles, are also provided in the popup box.

### Overview diagram of sequenced seed plants

The two major groups of seed plants are the angiosperms and gymnosperms. Angiosperms are by far the most diverse extant plant

group, comprising more than 350,000 known species assigned to more than 400 families in 64 orders. The gymnosperms are a much smaller group, containing only 1100 living species, which is comparable to some of the larger angiosperm genera (e.g., *Solanum*, *Acacia* and *Rhododendron*) in terms of species numbers. By the end of 2024, 1700 angiosperm species and 26 gymnosperm species had been sequenced.

While the cladogram diagrams in PubPlant focus at the individual species level, the overview diagram provides a view at the taxonomic family level using an embedded progress bar. The cladogram displays the phylogenetic position of each plant family while the progress bar depicts the number of sequenced species in that family (Figure 4). The overview diagram also includes families lacking any sequenced species thus far. This reveals that there are





many plant families with no sequenced species (e.g., 25 of 38 families in the order Caryophyllales).

## Use case: sequenced food crops

As a use case, we evaluated the current status of sequenced food crops. Economically important food crop plants have always been a major target for sequencing efforts because this provides information about the genes responsible for agronomically favorable traits. The enormous genome sizes of some of these crops, mainly due to polyploidy, the presence of repetitive DNA (Jackson et al., 2011) and very long introns (Xu et al., 2024), hindered progress until the advent of long-read sequencing (Pellicer and Leitch, 2020). For example, the sizes of the onion and broad bean genomes are 16 Gbp (Hao et al., 2023) and 13 Gbp (Jayakodi et al., 2023), respectively. But thanks to the advances mentioned earlier, almost all of the food crop species listed in the FAOSTAT database (FAOSTAT, 2024) have now been sequenced.

Further analysis of the most highly sequenced plant families (Poaceae, Fabaceae, Solanaceae, Rosaceae and Brassicaceae) indicated that these also contain the largest number of food crop species (Figure 5, listed in Table 2). Other families such as Salicaceae, which do not contain any food crops, also may have sequenced species, but this reflects the numerous poplar and willow species in that family, which are important to the timber industry. Similarly, many species of Orchidaceae have been sequenced (Figure 5) despite the presence of only one crop species (vanilla), due to the economic importance of many orchids as ornamental plants.

Only a few food crops remain to be sequenced (Table 3), including several culinary spices belonging to the Apiaceae (e.g. anise), leek (Amaryllidaceae), gooseberry and currants (Grossulariaceae). Notably, the blackcurrant (Grossulariaceae) is one of the most recently sequenced crop plants (Ziegler et al., 2024). Almost all of the food crops yet to be sequenced rank in the bottom half of the top 120 food crops ordered by production quantity (Table 3). Although all important food crop plants have now been sequenced, the remaining unsequenced species tend to be important in countries that contribute

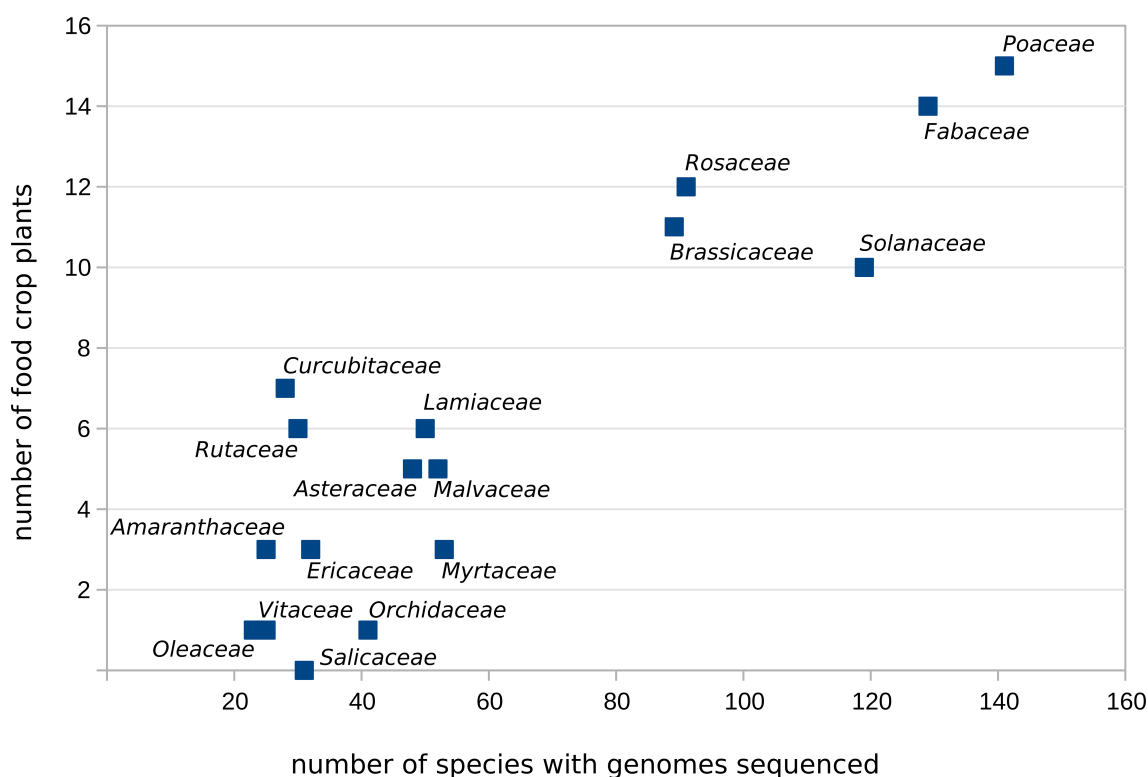


FIGURE 5

Scatter plot showing plant families in which more than 20 species have been sequenced. The x-axis indicates the number of species per family with sequenced genomes (evaluated February 2025) and the y-axis indicates the number of major food crops per family according to the FAOSTAT database (FAOSTAT, 2024), including those that have yet to be sequenced.

less to plant genome sequencing (Table 3). The countries that have contributed the most to the publication of crop plant genomes are China, the USA, Japan, Germany and Australia (Xie et al., 2024).

### Conclusion

In recent years, the publication of newly sequenced plant genomes has increased to such an extent that it has become a

weekly event (Figure 1). Periodic review articles provide a snapshot of the situation at the time the manuscripts were written but are quickly outdated. We therefore seek to highlight the online resource PubPlant, which provides up-to-date information on published plant genomes. PubPlant features a timeline view, where sequenced plant genomes are arranged in chronological order by the date of first publication (Figure 2), and a cladogram view, showing all sequenced plant species arranged according to their phylogeny (Figure 3). Separate cladograms are provided to display the phylogenetic

TABLE 2 Plant families with the highest numbers of sequenced genomes.

Taxonomic family	Plants with sequenced genomes <sup>a</sup>	Food crop species with sequenced genomes <sup>a,b</sup>
Poaceae	141 species	Sugar cane <i>Saccharum</i> x sp. Maize/corn <i>Zea mays</i> Wheat <i>Triticum aestivum</i> Rice <i>Oryza sativa</i> Barley <i>Hordeum vulgare</i> Sorghum <i>Sorghum bicolor</i> Millet <i>Setaria italica</i> , <i>Panicum miliaceum</i> , <i>Cenchrus americanus</i> , <i>Eleusine coracana</i> Oat <i>Avena sativa</i> Rye <i>Secale cereale</i> Fonio millet <i>Digitaria exilis</i>
Fabaceae	129 species	Soybean <i>Glycine max</i> Groundnut <i>Arachis hypogaea</i> Pea <i>Pisum sativum</i> Chickpea <i>Cicer arietinum</i> Cowpea <i>Vigna unguiculata</i> Broad bean <i>Vicia faba</i> Pigeon pea <i>Cajanus cajan</i> Lentil <i>Lens culinaris</i> Lupins <i>Lupinus albus</i> , <i>Lupinus angustifolius</i> Common (string) bean <i>Phaseolus vulgaris</i> Vetch <i>Vicia sativa</i> Bambara bean <i>Vigna subterranea</i> Locust bean <i>Ceratonia siliqua</i>
Rosaceae	91 species	Apple <i>Malus domestica</i> Peach and nectarine <i>Prunus persica</i> Pears <i>Pyrus communis</i> , <i>Pyrus pyrifolia</i> Plum and sloe <i>Prunus domestica</i> Strawberry <i>Fragaria</i> x <i>ananassa</i> Apricot <i>Prunus armeniaca</i> Almond <i>Prunus dulcis</i> Cherry <i>Prunus avium</i> Sour cherry <i>Prunus cerasus</i> Raspberry <i>Rubus idaeus</i> Quince <i>Cydonia oblonga</i>
Solanaceae	119 species	Potato <i>Solanum tuberosum</i> Tomato <i>Solanum lycopersicum</i> Eggplant <i>Solanum melongena</i> Chilies and peppers <i>Capsicum annuum</i> , <i>Capsicum frutescens</i> , <i>Capsicum chinense</i> , <i>Capsicum baccatum</i> , <i>Capsicum pubescens</i>
Brassicaceae	88 species	Cabbages <i>Brassica oleracea</i> , <i>Brassica rapa</i> Rapeseed <i>Brassica napus</i> Mustard <i>Sinapis alba</i> , <i>Brassica nigra</i> , <i>Brassica juncea</i> Other vegetables <i>Raphanus sativus</i> , <i>Eruca sativa</i> , <i>Lepidium sativum</i> , <i>Eutrema japonicum</i> , <i>Armoracia rusticana</i>

<sup>a</sup>Evaluated in February 2025.  
<sup>b</sup>Food crop categories according to FAOSTAT database.

TABLE 3 Food crop plants that have not been sequenced, or have been sequenced very recently.

Food crop category <sup>a</sup>	Ranking by production quantity <sup>b</sup>	Total production quantity (kiloton)	Main producers by production quantity (production share) <sup>c</sup>	Current status in genome sequencing <sup>d</sup>
Anise, badian, coriander, cumin, caraway, fennel, juniper berries <i>Pimpinella anisum</i> <i>Illicium verum</i> <i>Coriandrum sativum</i> <i>Cuminum cyminum</i> <i>Carum carvi</i> <i>Foeniculum vulgare</i> <i>Juniperus communis</i>	76th	2751	India (68.6%) Turkey (12.6%)	not sequenced (except <i>Coriandrum sativum</i> )
Leek <i>Allium ampeloprasum</i>	80th	2109	Indonesia (30.3%) Turkey (8.0%)	not sequenced
Yerba Maté <i>Ilex paraguariensis</i>	82nd	1653	Argentina (56.3%) Brazil (37.4%)	sequenced (Vignale et al., 2025)
Currants <i>Ribes nigrum</i> <i>Ribes rubrum</i>	97th	764	Russian Fed. (66.6%) Poland (19.1%)	<i>Ribes nigrum</i> sequenced (Ziegler et al., 2024)
Yautia <i>Xanthosoma sagittifolium</i>	103rd	396	Cuba (22.7%) Venezuela (22.3%)	not sequenced
Kola nuts <i>Cola acuminata</i> <i>Cola nitida</i>	104th	315	Nigeria (55.3%) Côte d'Ivoire (18.6%)	not sequenced
Gooseberry <i>Ribes uva-crispa</i>	112th	95	Russian Fed. (89.0%) Ukraine (8.0%)	not sequenced
Brazil nut <i>Bertholletia excelsa</i>	113th	79	Brazil (48.1%) Bolivia (42.9%)	sequenced (Wang et al., 2025)
Locust bean <i>Ceratonia siliqua</i>	114th	56	Turkey (44.5%) Morocco (39.1%)	sequenced (Akutsu et al., 2024)
Peppermint, spearmint <i>Mentha x piperita</i> <i>Mentha spicata</i>	115th	51	Morocco (84.0%) Argentina (13.7%)	<i>Mentha x piperita</i> Sequenced (Talbot et al., 2024)

<sup>a</sup>Food crop categories according to the FAOSTAT database.  
<sup>b</sup>Ranking among the top 120 food crop categories (data for 2022).  
<sup>c</sup>Top two countries with the largest share of world production (data for 2022).  
<sup>d</sup>Evaluated in February 2025.

positions of sequenced flowering and non-flowering plants. A summary diagram shows the phylogenetic position of all sequenced seed plants down to the taxonomic family rank, including those without any species sequenced thus far. It shows the total number of species for each family, the number of sequenced species, and the sequenced genera, while highlighting recently published genomes (Figure 4).

As a use case for PubPlant, we evaluated the status of sequenced food crop plants, which tend to be prioritized for genome sequencing. Unsurprisingly, almost all major food crop species (FAOSTAT food crop categories ranked by production quantity, full list in Supplementary Table 2) have already been sequenced. In

addition, the five plant families with the greatest number of sequenced plant genomes (Poaceae, Fabaceae, Rosaceae, Solanaceae and Brassicaceae) are also those containing the most major food crop species (Table 2).

PubPlant has been available online for more than 9 years and has been widely used, including as a resource to prepare review articles summarizing plant genome sequencing progress (Jiao and Schneeberger, 2017; Hao et al., 2022; Bernal-Gallardo and de Folter, 2024) and to evaluate the status of sequenced medicinal plants (Cheng et al., 2021b). Whereas other plant genome resources release irregular updates, PubPlant is updated on a monthly basis, providing a simple and intuitive source of current and historical



information on sequenced plant genomes for the benefit of the entire plant research community.

## Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## Author contributions

RS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing, Resources. MB: Conceptualization, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing, Funding acquisition. BU: Project administration, Supervision, Writing – original draft, Writing – review & editing, Funding acquisition.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. RS, MB and BU are supported by the German Federal Ministry of Research, Technology and Space (BMFTR) in the frame of the German Network for Bioinformatics Infrastructure (de.NBI).

## References

- Akutsu, M., Shinozawa, A., Nishiyama, T., Sakata, Y., and Hiwatashi, Y. (2024). *De novo* sequencing allows genome-wide identification of genes involved in galactomannan synthesis in locust bean (*Ceratonia siliqua*). *DNA Res.* 31, dsae033. doi: 10.1093/dnares/dsae033
- COL. (2025). *Catalogue of life*. (Amsterdam, The Netherlands). Available online at: <https://www.catalogueoflife.org> (Accessed February 10, 2025).
- Bechteler, J., Peñaloza-Bojacá, G., Bell, D., Gordon Burleigh, J., McDaniel, S. F., Christine Davis, E., et al. (2023). Comprehensive phylogenomic time tree of bryophytes reveals deep relationships and uncovers gene incongruences in the last 500 million years of diversification. *Am. J. Bot.* 110, e16249. doi: 10.1002/ajb2.v110.11
- Bernal-Gallardo, J. J., and de Folter, S. (2024). Plant genome information facilitates plant functional genomics. *Planta* 259, 117. doi: 10.1007/s00425-024-04397-z
- Bozan, I., Achakagari, S. R., Anglin, N. L., Ellis, D., Tai, H. H., and Strömvik, M. V. (2023). Pan-genome analyses reveal impact of transposable elements and ploidy on the evolution of potato species. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2211117120. doi: 10.1073/pnas.2211117120
- Chen, F., Song, Y., Li, X., Chen, J., Mo, L., Zhang, X., et al. (2019). Genome sequences of horticultural plants: past, present, and future. *Hortic. Res.* 6, 112. doi: 10.1038/s41438-019-0195-6
- Chen, S., Wang, P., Kong, W., Chai, K., Zhang, S., Yu, J., et al. (2023). Gene mining and genomics-assisted breeding empowered by the pangenome of tea plant *Camellia sinensis*. *Nat. Plants* 9, 1986–1999. doi: 10.1038/s41477-023-01565-z
- Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., and Li, H. (2021a). Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nat. Methods* 18, 170–175. doi: 10.1038/s41592-020-01056-5
- Cheng, Q. Q., Ouyang, Y., Tang, Z. Y., Lao, C. C., Zhang, Y. Y., Cheng, C. S., et al. (2021b). Review on the development and applications of medicinal plant genomes. *Front. Plant Sci.* 12, 791219. doi: 10.3389/fpls.2021.791219
- Clarivate (2025). *Web of Science database*. Available online at: <https://www.webofscience.com> (Accessed February 10, 2024).
- Dumschott, K., Schmidt, M. H.-W., Chawla, H. S., Snowdon, R., and Usadel, B. (2020). Oxford Nanopore Sequencing: New opportunities for plant genomics? *J. Exp. Bot.* 71, 5313–5322. doi: 10.1093/jxb/eraa263
- FAOSTAT (2024). *Food and agriculture organization of the united nations*. Available online at: <https://www.fao.org/faostat> (Accessed March 9, 2024).
- Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., et al. (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* 51, 1044–1051. doi: 10.1038/s41588-019-0410-2
- Golicz, A. A., Batley, J., and Edwards, D. (2016). Towards plant pangenomics. *Plant Biotechnol. J.* 14, 1099–1105. doi: 10.1111/pbi.2016.14.issue-4
- Govaerts, R., Nic Lughadha, E., Black, N., Turner, R., and Paton, A. (2021). The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. *Sci. Data* 8, 215. doi: 10.1038/s41597-021-00997-6
- Guiry, M. D., and Guiry, G. M. (2025). *AlgaeBase* (University of Galway, Galway, Ireland). Available at: <https://www.algaebase.org> (Accessed March 09, 2025).
- Hao, F., Liu, X., Zhou, B., Tian, Z., Zhou, L., Zong, H., et al. (2023). Chromosome-level genomes of three key *Allium* crops and their trait evolution. *Nat. Genet.* 55, 1976–1986. doi: 10.1038/s41588-023-01546-0
- Hao, L., Xu, W., Qi, G., Xin, T., Xu, Z., Lei, H., et al. (2022). GAGE is a method for identification of plant species based on whole genome analysis and genome editing. *Commun. Biol.* 5, 947. doi: 10.1038/s42003-022-03894-9
- Hassler, M. (2025). “Checklist of ferns and lycophytes of the world,” in *Catalogue of life*. Eds. Bánki, O., Roskov, Y., Döring, M., Ower, G., Hernández Robles, D. R., Plata Corredor, C. A., et al. (Amsterdam, Netherlands: Catalogue of life partners).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2025.1603547/full#supplementary-material>

### SUPPLEMENTARY TABLE 1

List of published sequenced plant genomes in chronological order.

### SUPPLEMENTARY TABLE 2

List of food crop categories (FAOSTAT) ranked by production quantity.

- Hufford, M. B., Seetharam, A. S., Woodhouse, M. R., Chougule, K. M., Ou, S., Liu, J., et al. (2021). *De novo* assembly, annotation, and comparative analysis of 26 diverse maize genomes. *Science* 373, 655–662. doi: 10.1126/science.abg5289
- Jackson, S. A., Iwata, A., Lee, S.-H., Schmutz, J., and Shoemaker, R. (2011). Sequencing crop genomes: approaches and applications: Tansley review. *New Phytol.* 191, 915–925. doi: 10.1111/j.1469-8137.2011.03804.x
- Jayakodi, M., Golicz, A. A., Kreplak, J., Fechete, L. I., Angra, D., Bednář, P., et al. (2023). The giant diploid faba genome unlocks variation in a global protein crop. *Nature* 615, 652–659. doi: 10.1038/s41586-023-05791-5
- Jayakodi, M., Padmarasu, S., Haberer, G., Bonthala, V. S., Gundlach, H., Monat, C., et al. (2020). The barley pan-genome reveals the hidden legacy of mutation breeding. *Nature* 588, 284–289. doi: 10.1038/s41586-020-2947-8
- Jiao, W.-B., and Schneeberger, K. (2017). The impact of third generation genomic technologies on plant genome assembly. *Curr. Opin. Plant Biol.* 36, 64–70. doi: 10.1016/j.pbi.2017.02.002
- Jiao, C., Xie, X., Hao, C., Chen, L., Xie, Y., Garg, V., et al. (2024). Pan-genome bridges wheat structural variations with habitat and breeding. *Nature* 637, 384–393. doi: 10.1038/s41586-024-08277-0
- Kersey, P. J. (2019). Plant genome sequences: past, present, future. *Curr. Opin. Plant Biol.* 48, 1–8. doi: 10.1016/j.pbi.2018.11.001
- Khan, A. W., Garg, V., Roorkiwal, M., Golicz, A. A., Edwards, D., and Varshney, R. K. (2020). Super-pan genome by integrating the wild side of a species for accelerated crop improvement. *Trends Plant Sci.* 25, 148–158. doi: 10.1016/j.tplants.2019.10.012
- Kress, W. J., Soltis, D. E., Kersey, P. J., Wegrzyn, J. L., Leebens-Mack, J. H., Gostel, M. R., et al. (2022). Green plant genomes: What we know in an era of rapidly expanding opportunities. *Proc. Natl. Acad. Sci. U.S.A.* 119, e2115640118. doi: 10.1073/pnas.2115640118
- Michael, T. P., and Jackson, S. (2013). The first 50 plant genomes. *Plant Genome* 6, eplantgenome2013.03.0001in. doi: 10.3835/plantgenome2013.03.0001in
- Pellicer, J., and Leitch, I. J. (2020). The Plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. *New Phytol.* 226, 301–305. doi: 10.1111/nph.v226.2
- Plants of the World Online (POWO). (2025). *Facilitated by the royal botanic gardens* (Kew, London, UK). Available at: <https://powo.science.kew.org/>.
- PubMed (2025). (Bethesda (MD: National Library of Medicine (US). (Accessed February 15, 2025).
- Pucker, B., Irisarri, I., de Vries, J., and Xu, B. (2022). Plant genome sequence assembly in the era of long reads: Progress, challenges and future directions. *Quant. Plant Biol.* 3, e5. doi: 10.1017/qpb.2021.18
- Qin, P., Lu, H., Du, H., Wang, H., Chen, W., Chen, Z., et al. (2021). Pan-genome analysis of 33 genetically diverse rice accessions reveals hidden genomic variations. *Cell* 184, 3542–3558. doi: 10.1016/j.cell.2021.04.046
- Rautiainen, M., Nurk, S., Walenz, B. P., Logsdon, G. A., Porubsky, D., Rhie, A., et al. (2023). Telomere-to-telomere assembly of diploid chromosomes with Verkko. *Nat. Biotechnol.* 41, 1474–1482. doi: 10.1038/s41587-023-01662-6
- Schoch, C. L., Ciufu, S., Domrachev, M., Hottton, C. L., Kannan, S., Khovanskaya, R., et al. (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database (Oxford)* 2020, baaa062. doi: 10.1093/database/baaa062
- Shirasawa, K., Harada, D., Hirakawa, H., Isobe, S., and Kole, C. (2021). Chromosome-level *de novo* genome assemblies of over 100 plant species. *Breed. Sci.* 71, 117–124. doi: 10.1270/jsbbs.20146
- Sun, Y., Shang, L., Zhu, Q.-H., Fan, L., and Guo, L. (2022). Twenty years of plant genome sequencing: achievements and challenges. *Trends Plant Sci.* 27, 391–401. doi: 10.1016/j.tplants.2021.10.006
- Talbot, S. C., Pandelova, I., Lange, B. M., and Vining, K. J. (2024). A first look at the genome structure of hexaploid “Mitcham” peppermint (*Mentha × piperita* L.). *G3 (Bethesda)* 14, jkae195. doi: 10.1093/g3journal/jkae195
- Tang, D., Jia, Y., Zhang, J., Li, H., Cheng, L., Wang, P., et al. (2022). Genome evolution and diversity of wild and cultivated potatoes. *Nature* 606, 535–541. doi: 10.1038/s41586-022-04822-x
- Tariq, A., Meng, M., Jiang, X., Bolger, A., Beier, S., Buchmann, J. P., et al. (2024). In-depth exploration of the genomic diversity in tea varieties based on a newly constructed pangenome of *Camellia sinensis*. *Plant J.* 119, 2096–2115. doi: 10.1111/tpj.v119.4
- The Angiosperm Phylogeny Group. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* 181, 1–20. doi: 10.1111/boj.2016.181.issue-1
- The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815. doi: 10.1038/35048692
- The Pteridophyte Phylogeny Group. (2016). A community-derived classification for extant lycophytes and ferns: PPG I. *J. Syst. Evol.* 54, 563–603. doi: 10.1111/jse.12229
- van Rens, W. M. J., Schmidt, M. H.-W., Effgen, S., Le, D. B., Wang, Y., Zaidan, M. W. A. M., et al. (2022). A chromosome scale tomato genome built from complementary PacBio and Nanopore sequences alone reveals extensive linkage drag during breeding. *Plant J.* 110, 572–588. doi: 10.1111/tpj.v110.2
- Vignale, F. A., Hernandez Garcia, A., Modenutti, C. P., Sosa, E. J., Defelipe, L. A., Oliveira, R., et al. (2025). Yerba mate (*Ilex paraguariensis*) genome provides new insights into convergent evolution of caffeine biosynthesis. *Elife* 14, e104759. doi: 10.7554/eLife.104759
- Wang, H., Caetano-Andrade, V., Boivin, N., Clement, C. R., Ayala, W. E., Melinski, R. D., et al. (2025). Long-term human influence on the demography and genetic diversity of the hyperdominant *Bertholletia excelsa* in the Amazon Basin. *Curr. Biol.* 35, 629–639. doi: 10.1016/j.cub.2024.12.023
- Xie, L., Gong, X., Yang, K., Huang, Y., Zhang, S., Shen, L., et al. (2024). Technology-enabled great leap in deciphering plant genomes. *Nat. Plants* 10, 551–566. doi: 10.1038/s41477-024-01655-6
- Xu, S., Chen, R., Zhang, X., Wu, Y., Yang, L., Sun, Z., et al. (2024). The evolutionary tale of lilies: Giant genomes derived from transposon insertions and polyploidization. *Innovation (Camb.)* 5, 100726. doi: 10.1016/j.xinn.2024.100726
- Yang, Y., Ferguson, D. K., Liu, B., Mao, K. S., Gao, L. M., Zhang, S. Z., et al. (2022). Recent advances on phylogenomics of gymnosperms and a new classification. *Plant Divers.* 44, 340–350. doi: 10.1016/j.pld.2022.05.003
- Zhou, Y., Zhang, Z., Bao, Z., Li, H., Lyu, Y., Zan, Y., et al. (2022). Graph pangenome captures missing heritability and empowers tomato breeding. *Nature* 606, 527–534. doi: 10.1038/s41586-022-04808-9
- Ziegler, F. M. R., Rosenthal, V., Vallarino, J. G., Genzel, F., Spettmann, S., Seliga, L., et al. (2024). A full genome assembly reveals drought stress effects on gene expression and metabolite profiles in blackcurrant (*Ribes nigrum* L.). *Hortic. Res.* 12, uhac313. doi: 10.1093/hr/uhac313
- Zuntini, A. R., Carruthers, T., Maurin, O., Bailey, P. C., Leempoel, K., Brewer, G. E., et al. (2024). Phylogenomics and the rise of the angiosperms. *Nature* 629, 843–850. doi: 10.1038/s41586-024-07324-0